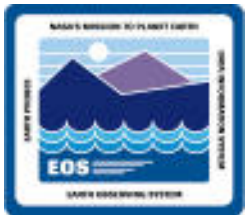
Agenda

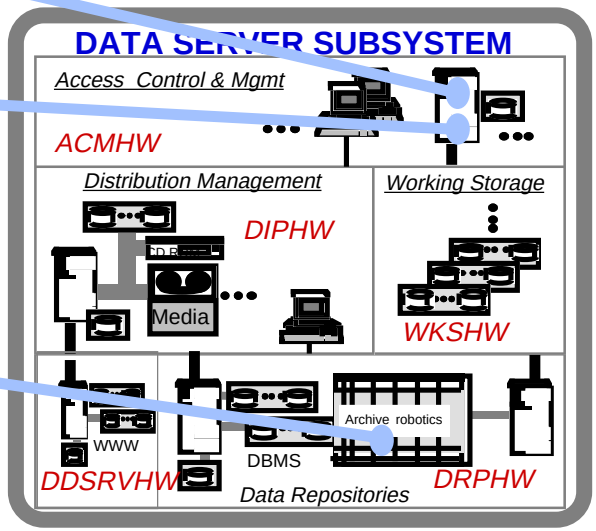


- Release B Capabilities
- Software CSCs
- Software CSC Status
- Progress Since IDR-B
- ESDT Design
- Design Trades
- Sybase/Illustra Benchmark Status
- High Level HW/SW Physical Design
- Detailed HW/SW Architecture
- Transition Topics
- Next Steps
- Design Topics

High Level HW/SW Physical Design

Three Processes

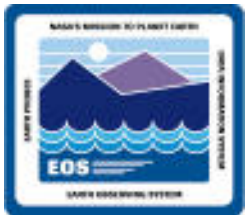
- **ScienceDataServer**
  - Process supporting the access to the advertised data type services as well as the implementations of those services.
- **SDSRVAdmin**
  - Supports the SDSRV Administrators and Operators. Primarily access to configuration of the SDSRV and monitor and control of SDSRV processing.
- **COTS Database Engine (Illustra)**
  - Commercial off-the-shelf ORDBMS used for storing and searching metadata.



One Library

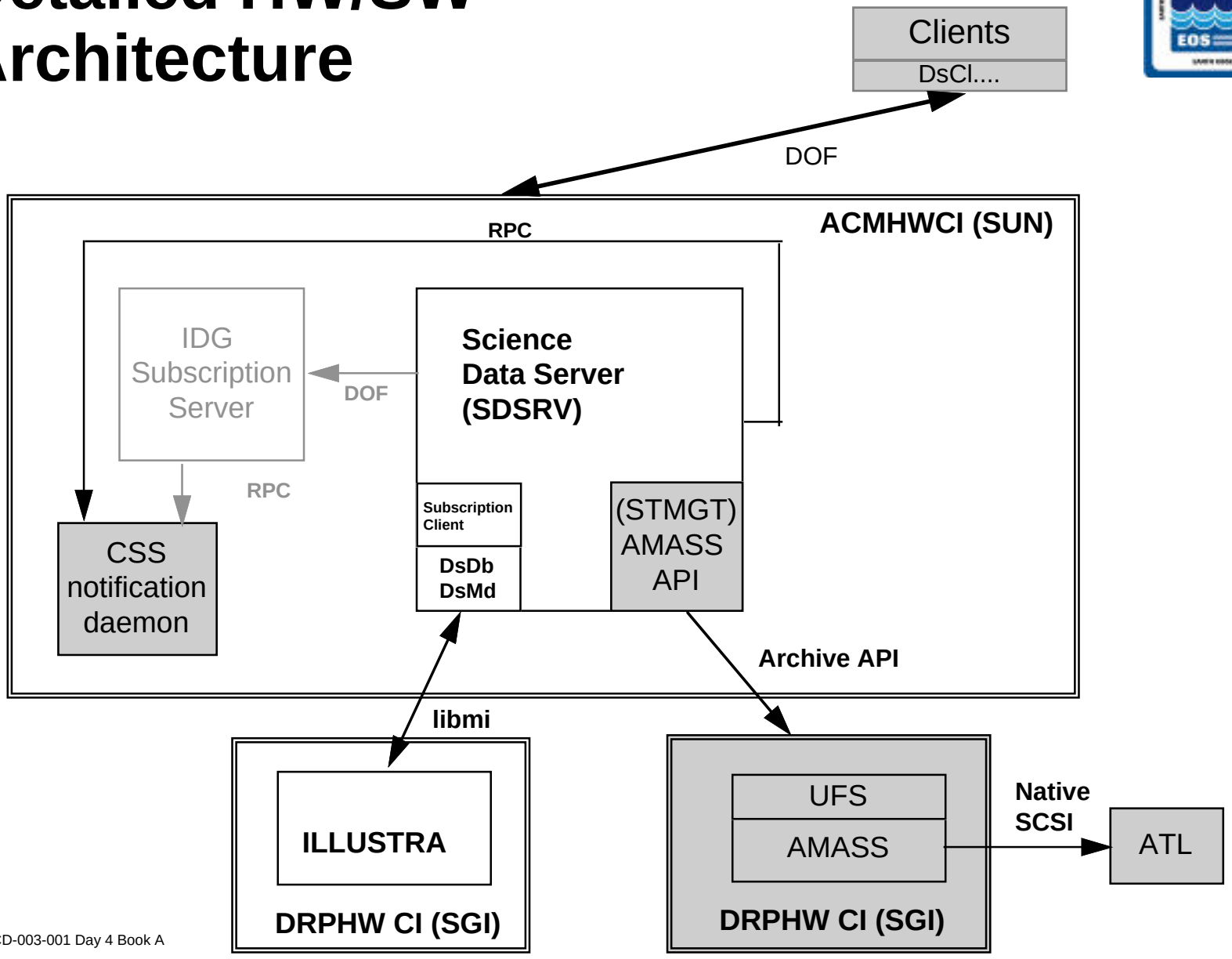
- **SDSRVClient Library**
  - Software toolkit library for use by SDSRV Client applications for connecting to and interacting with the SDSRV.

Agenda

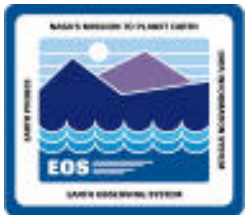


- Release B Capabilities
- Software CSCs
- Software CSC Status
- Progress Since IDR-B
- ESDT Design
- Design Trades
- Sybase/Illustra Benchmark Status
- High Level HW/SW Physical Design
- Detailed HW/SW Architecture
- Transition Topics
- Next Steps
- Design Topics

Detailed HW/SW Architecture

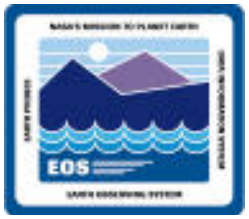


Agenda



- Release B Capabilities
- Software CSCs
- Software CSC Status
- Progress Since IDR-B
- ESDT Design
- Design Trades
- Sybase/Illustra Benchmark Status
- High Level HW/SW Physical Design
- Detailed HW/SW Architecture
- Transition Topics
- Next Steps
- Design Topics

Transition Assumptions



There is no requirement to run Release A and B SDSRV DBMS Servers in parallel.

Based on Release A DAAC Specific Design Documents, each DAAC will have less than 1GB of metadata to be loaded from Sybase into Illustra.

Migration of Release B V0 data into Illustra during the Release B timeframe is considered a separate transition.

There will be few data type changes or field length changes between the Sybase and Illustra schemas with little “data cleanup” required as in migrations from legacy systems to new systems.

Transition Approach

The transition from Sybase to Illustra will be handled as an "initial database load," where at the start of Release B, the Sybase DBMS is turned off and the Illustra DBMS is turned on.

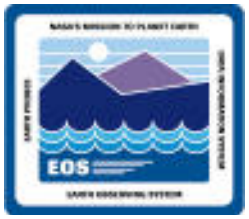
Stage 1:

- **Construct flat files from Sybase that contain all of the data as of a certain timestamp**
  - **map Sybase and Illustra schemas**
  - **develop software tooread Sybase and generate flat files that correlate to Illustra tables**
- **Load data into Illustra using "direct access method"**
  - **load independent tables in parallel**
  - **minimize "recovery" time if Illustra goes down during load**

Stage 2:

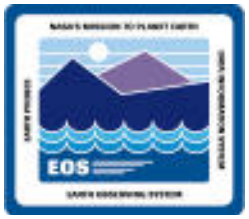
- **Construct flat files from Sybase that cover the delta between the initial capture and the cut-off point for the Release B switch**
- **Load data into Illustra using approach from Stage 1**
- **Validate that all data from Sybase has been loaded into Illustra**

Agenda



- Release B Capabilities
- Software CSCs
- Software CSC Status
- Progress Since IDR-B
- ESDT Design
- Design Trades
- Sybase/Illustra Benchmark Status
- High Level HW/SW Physical Design
- Detailed HW/SW Architecture
- Transition Topics
- Next Steps
- Design Topics

Next Steps



Continue Resolution of Design Issues and Topics

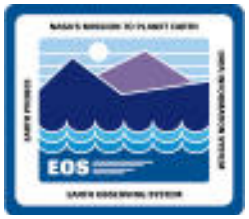
- **Data Server Detailed Design Meetings (DDM)**

Complete Detailed Design

- **Release A Retrofit**
- **Incorporate Updates to Object Model**
- **Complete Sybase/Illustra Benchmark Prototype**
- **Complete Schema Design and Update 311-CD-002-005**
- **Conduct Design Inspections**
- **Republish 305-CD-024-002**

Conduct Delta Detailed Design Review (3DR)

Agenda



- Release B Capabilities
- Software CSCs
- Software CSC Status
- Progress Since IDR-B
- ESDT Design
- Design Trades
- Sybase/Illustra Benchmark Status
- High Level HW/SW Physical Design
- Detailed HW/SW Architecture
- Transition Topics
- Next Steps
- Design Topics

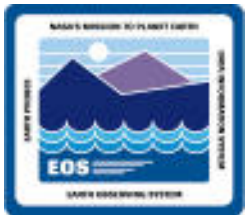
IDR-B Issues

Issue, Comment, Question, AI, or RID		Code	Requester	Recorder	Comment
2	Want to be able to define data quality level in a subscription.	I	Lyn Oleson	Craig Schillhahn	3DR
6	What is going to be in the metadata? (what is expected to be in the metadata by ECS.	BC	Lyn Oleson	Craig Schillhahn	Completed In DID 311
6	Provide URL for ESDT definitions.	AI	Lyn Oleson	Mike Burnett	Complete
33	How do we group or specify which ESDTs are needed for a specific data product?	I	Lyn Oleson	Craig Schillhahn	The concept of related granule information being packaged with an ESDT granule has been addressed by the granule package. Granule packages are being defined for each ESDT.
34	Where does the doc/presentation discuss the population of Lim/DIM with "granual packages". How can the pieces of a granual package be tracked?	I	Lyn Oleson	Craig Schillhahn	Submitted RID #11. Awaiting Approval
51	Emphasize Update loading in Illustra prototyping. Should interact with DAAC experience over performance expectations. (Lyn offered data.)	I	Lyn Oleson	Steve Marley	Present At CDR
53	Dataserver at the OPS workshop needs to discuss the list of "scripted" subscription actions in response to events that ECS provides upon delivery, and what DAACs are expected to write.	I	Steve Marley	Craig Schillhahn	Conducted Ops Workshop

IDR-B Promises

Number	SDSRV IDR Promises
7	Stability of the Initial ESDTs
7	Operational and Developer Document on ESDTs
8	Establish Application Domain on Session Management
9	Validation of Data Type Service Matrix
10	Bundling of Notifications of Subscriptions
14	Finalize ASF Requirements and Refine the Design
14	Analyze Planning Implications - Nested Production Dependencies
14	Analyze Impact of Run Time Parameters
15	Analyze Feedback from Release A Implementation

RID #5 Meeting Actions

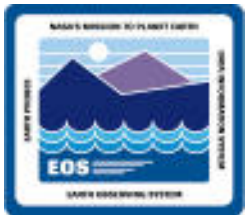


Action Item	Title	Description	Assignment	Contact	Due Date	Comments
4	CDR Functionality Demonstrations	ECS Should include a scenario showing how PDPS, DSS and Client work together. ECS should also present detailed scenarios of virtual product concept using ASTER as an example.	ECS	SDSRV	6-Jun	
5	DAR ID from Japan	Is the DAR ID actually in the metadata?	ECS	SDSRV	Update 4/5	Our design for ASTER DAR ID is to have a product specific metadata attribute for ASTER products. Data Modeling is responsible for identifying product specific metadata for all products.
6	Detailed Subscription Scenario	Provide detailed scenario for how the subscriptions work, e.g., a) how to subscribe to metadata update service, b) belated discovery of corrupted data, and c) manual process of looking up subscribers.	ECS	SDSRV	6-Jun	
7	Notification of Users When Products Change	How to look up all subscriptions going to folks who are PGE-related and all folks who are just science users.	ECS	SDSRV	6-Jun	
8	Different Classes for Data Access	Lyn Oleson suggested DSS might want to consider different classes of users. Need for a flag for data validation and verification prior to access by users?	ECS	SDSRV	6-Jun	
10	Coordination for Data and Browse Ingest	Please provide scenarios that show realistically how metadata and data arrive over time.	ECS	SDSRV	6-Jun	
11	Feedback from Community About "Browse" Definition	Science community has been unable to completely define what and if a browse product is for each of the data products.	ESDIS	Ben Kobler	5-Apr	Ted Meyer is to prepare a paper identifying browse data. ECS is proceeding with our assumption on browse products with 1Mbyte size. DSS doesn't deal with content of package only sends and receives the package. Continues to be open for ESDIS.
14	Temporary URs	A scenario should be provided describing the steps in a subset request showing the transformation state of the granule.	ECS/CDR	SDSRV	6-Jun	
17	Coordination for Metadata Problem Reporting	We need to be sure that CIDM knows how to handle the DSS metadata problem report and that there is no confusion with Trouble Tickets. Need to assess the role of metadata problem report in the system context.	ECS	SDSRV	6-Jun	

RID #5 Meeting Actions (cont.)

Action Item	Title	Description	Assignment	Contact	Due Date	Comments
18	Description of DAR Requirements	What subsystem at CDR will describe the DAR scenario? Associated with the DAR, what "up front" capabilities will be available in the client GUI tool?	CDR	SDSRV	6-Jun	
19	How Algorithms are Made Available	A scenario for acquiring the Science SW Algorithm Programs from the archive should be included in the documentation. Includes several issues.	ECS	SDSRV	6-Jun	
20	Description of How New Product Versions are Handled	State explicitly what versioning is and what the requirements are for it. Show how the system is designed to handle it.	CDR	SDSRV	6-Jun	
23	How Many Data Granules Represented in DB Benchmark?	How is the Benchmark going to simulate the varying data types representative of Releases A and B? Please include the number of each different type in the study (how many rows per group?)	CDR	SDSRV	5-Apr	Approximately 46 Million
24	How Does Subscription Processing Impact Database Performance?	There was a question about the distribution of the 100 users being simulated. Several suggestions were made.	ECS	SDSRV	5-Apr	There is not a subscription simulator but the subscription load could be considered part of the 100 concurrent users. SDSRV doesn't distinguish between actions being fulfilled as actual requests vs. on behalf of clients from the subscription server.
26	How do Priorities In Data server Allow Management of Limited Resources?	Chris Harris requested that we determine what the dynamics of changing priorities are and how many different types of classes can you have at a given time? How many flavors of users are there?	ECS	SDSRV	6-Jun	
27	How is Priority Information in DARs and in On-Demand Processing Passed to Processing?	DAR processing scenarios were discussed; a controversial point concerned how the user profile was passed to the SDSRV. Oleson stressed how different levels of users may submit a DAR.	ECS	SDSRV	6-Jun	
29	QA for On-Demand Processing and Re-Generation of Higher Level Products After Data Restoration	Check on QA plans both in regular processing and in the on-demand processing.	ECS	SDSRV	6-Jun	

Ops Workshop Action Items



SDSRV Action Items Number	Title	Description	Status	Due Date	Document Trace
Action Item #9	Verify Metadata	Need clarification on "verify metadata" step. Are there unnecessary redundancies?	Closed		
Action Item #10	View of Inventory	Desire to control the view of the inventory on a granule/file level. Explore a public vs. private (restricted) advertising of data availability as a means of restricting access to a targeted granule/file.	Open		
Action Item #11	EOS Products	Need to identify file dependencies on EOS products & VO migrated data.	Open		
Action Item #12	Partial Failure	Is requestor notified of a partial failure of a request?	For store and retrieval requests: No, there are no partial successes for store or retrieval operations.		
Action Item #51	Submitted Requests	For submitted requests by User Services on a user's behalf, user services will let the user know how big their request is.	Reject -User Services		
Action Item #54	Sybase to Illustra Switchover	What will happen if during the switchover from Sybase to Illustra, it is discovered at a later date that the Illustra DB is corrupted to the point that reverting back to Sybase without lost data is not possible?	Accept		
Action Item #55	Sybase to Illustra Switchover	Consider the number of granules as well as volume for the Sybase to Illustra and VO transition include in the white paper.	Open		
Action Item #57	Fix Scenario Disconnects	Fix any disconnects between this scenario and the subscription event error/resolution scenarios.	Open		
Action Item #73	Order processing Cancellation	At what point in processing is it too late to cancel an order?	Closed		
Action Item #78	Data Access Restrictions	Are data access restrictions reflected in the metadata model? If not, they should be.	Reject		
Action Item #85	Group Subscriptions	Need a scenario to show how group subscriptions will be handled. (How do you do a batch notify of a problem?)	Rejected - IDG Function		